



SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: Chen, J. Don
Li, Hui
- (ii) TITLE OF INVENTION: Transcriptional Coactivator for Nuclear
Hormone Receptors
- (iii) NUMBER OF SEQUENCES: 2
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: Lahive and Cockfield
 - (B) STREET: 28 State Street
 - (C) CITY: Boston
 - (D) STATE: MA
 - (E) COUNTRY: USA
 - (F) ZIP: 02109
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Floppy disk
 - (B) COMPUTER: IBM PC compatible
 - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (D) SOFTWARE: PatentIn Release #1.0, Version #1.25
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER:
 - (B) FILING DATE:
 - (C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: Liepmann, W. Hugo
 - (B) REGISTRATION NUMBER: 20,407
 - (C) REFERENCE/DOCKET NUMBER: UMM-026-1
- (ix) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: 617-227-7400
 - (B) TELEFAX: 617-742-4214

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 4496 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (ix) FEATURE:

(A) NAME/KEY: CDS
(B) LOCATION: 86..4338

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GCTGGATGGT GGACTCAGAG ACCAATAAAA ATAACTGCT TGAACATCCT TTGACTGGTT	60
AGCCAGTTGC TGATGTATAT TCAAG ATG AGT GGA TTA GGA GAA AAC TTG GAT	112
Met Ser Gly Leu Gly Glu Asn Leu Asp	
1 5	
CCA CTG GCC AGT GAT TCA CGA AAA CGC AAA TTG CCA TGT GAT ACT CCA	160
Pro Leu Ala Ser Asp Ser Arg Lys Arg Lys Leu Pro Cys Asp Thr Pro	
10 15 20 25	
GGA CAA GGT CTT ACC TGC AGT GGT GAA AAA CGG AGA CGG GAG CAG GAA	208
Gly Gln Gly Leu Thr Cys Ser Gly Glu Lys Arg Arg Arg Glu Gln Glu	
30 35 40	
AGT AAA TAT ATT GAA GAA TTG GCT GAG CTG ATA TCT GCC AAT CTT AGT	256
Ser Lys Tyr Ile Glu Glu Leu Ala Glu Leu Ile Ser Ala Asn Leu Ser	
45 50 55	
GAT ATT GAC AAT TTC AAT GTC AAA CCA GAT AAA TGT GCG ATT TTA AAG	304
Asp Ile Asp Asn Phe Asn Val Lys Pro Asp Lys Cys Ala Ile Leu Lys	
60 65 70	
GAA ACA GTA AGA CAG ATA CGT CAA ATA AAA GAG CAA GGA AAA ACT ATT	352
Glu Thr Val Arg Gln Ile Arg Gln Ile Lys Glu Gln Gly Lys Thr Ile	
75 80 85	
TCC AAT GAT GAT GAT GTT CAA AAA GCC GAT GTA TCT TCT ACA GGG CAG	400
Ser Asn Asp Asp Asp Val Gln Lys Ala Asp Val Ser Ser Thr Gly Gln	
90 95 100 105	
GGA GTT ATT GAT AAA GAC TCC TTA GGA CCG CTT TTA CTT CAG GCA TTG	448
Gly Val Ile Asp Lys Asp Ser Leu Gly Pro Leu Leu Leu Gln Ala Leu	
110 115 120	
GAT GGT TTC CTA TTT GTG GTG AAT CGA GAG GCA AAC ATT GTA TTT GTA	496
Asp Gly Phe Leu Phe Val Val Asn Arg Glu Ala Asn Ile Val Phe Val	
125 130 135	
TCA GAA AAT GTC ACA CAA TAC CTG CAA TAT AAG CAA GAG GAC CTG GTT	544
Ser Glu Asn Val Thr Gln Tyr Leu Gln Tyr Lys Gln Glu Asp Leu Val	
140 145 150	
AAC ACA AGT GTT TAC AAT ATC TTA CAT GAA GAA GAC AGA AAG GAT TTT	592
Asn Thr Ser Val Tyr Asn Ile Leu His Glu Glu Asp Arg Lys Asp Phe	
155 160 165	
CTT AAG AAT TTA CCA AAA TCT ACA GTT AAT GGA GTT TCC TGG ACA AAT	640
Leu Lys Asn Leu Pro Lys Ser Thr Val Asn Gly Val Ser Trp Thr Asn	

170	175	180	185	
GAG CCC CAA AGA CAA AAA AGC CAT ACA TTT AAT TGC CGT ATG TTG ATG				688
Glu Pro Gln Arg Gln Lys Ser His Thr Phe Asn Cys Arg Met Leu Met				
	190	195	200	
AAA ACA CCA CAT GAT ATT CTG GAA GAC ATA AAC GCC AGT CCT GAA ATG				736
Lys Thr Pro His Asp Ile Leu Glu Asp Ile Asn Ala Ser Pro Glu Met				
	205	210	215	
CGC CAG AGA TAT GAA ACA ATG CAG TGC TTT GCC CTG TCT CAG CCA CGA				784
Arg Gln Arg Tyr Glu Thr Met Gln Cys Phe Ala Leu Ser Gln Pro Arg				
	220	225	230	
GCT ATG ATG GAG GAA GGG GAA GAT TTG CAA TCT TGT ATG ATC TGT GTG				832
Ala Met Met Glu Glu Gly Glu Asp Leu Gln Ser Cys Met Ile Cys Val				
	235	240	245	
GCA CGC CGC ATT ACT ACA GGA GAA AGA ACA TTT CCA TCA AAC CCT GAG				880
Ala Arg Arg Ile Thr Thr Gly Glu Arg Thr Phe Pro Ser Asn Pro Glu				
	250	255	260	265
AGC TTT ATT ACC AGA CAT GAT CTT TCA GGA AAG GTT GTC AAT ATA GAT				928
Ser Phe Ile Thr Arg His Asp Leu Ser Gly Lys Val Val Asn Ile Asp				
	270	275	280	
ACA AAT TCA CTG AGA TCC TCC ATG AGG CCT GGC TTT GAA GAT ATA ATC				976
Thr Asn Ser Leu Arg Ser Ser Met Arg Pro Gly Phe Glu Asp Ile Ile				
	285	290	295	
CGA AGG TGT ATT CAG AGA TTT TTT AGT CTA AAT GAT GGG CAG TCA TGG				1024
Arg Arg Cys Ile Gln Arg Phe Phe Ser Leu Asn Asp Gly Gln Ser Trp				
	300	305	310	
TCC CAG AAA CGT CAC TAT CAA GAA GCT TAT CTT AAT GGC CAT GCA GAA				1072
Ser Gln Lys Arg His Tyr Gln Glu Ala Tyr Leu Asn Gly His Ala Glu				
	315	320	325	
ACC CCA GTA TAT CGA TTC TCG TTG GCT GAT GGA ACT ATA GTG ACT GCA				1120
Thr Pro Val Tyr Arg Phe Ser Leu Ala Asp Gly Thr Ile Val Thr Ala				
	330	335	340	345
CAG ACA AAA AGC AAA CTC TTC CGA AAT CCT GTA ACA AAT GAT CGA CAT				1168
Gln Thr Lys Ser Lys Leu Phe Arg Asn Pro Val Thr Asn Asp Arg His				
	350	355	360	
GGC TTT GTC TCA ACC CAC TTC CTT CAG AGA GAA CAG AAT GGA TAT AGA				1216
Gly Phe Val Ser Thr His Phe Leu Gln Arg Glu Gln Asn Gly Tyr Arg				
	365	370	375	
CCA AAC CCA AAT CCT GTT GGA CAA GGG ATT AGA CCA CCT ATG GCT GGA				1264
Pro Asn Pro Asn Pro Val Gly Gln Gly Ile Arg Pro Pro Met Ala Gly				
	380	385	390	

TGC AAC AGT TCG GTA GGC GGC ATG AGT ATG TCG CCA AAC CAA GGC TTA	1312
Cys Asn Ser Ser Val Gly Gly Met Ser Met Ser Pro Asn Gln Gly Leu	
395 400 405	
CAG ATG CCG AGC AGC AGG GCC TAT GGC TTG GCA GAC CCT AGC ACC ACA	1360
Gln Met Pro Ser Ser Arg Ala Tyr Gly Leu Ala Asp Pro Ser Thr Thr	
410 415 420 425	
GGG CAG ATG AGT GGA GCT AGG TAT GGG GGT TCC AGT AAC ATA GCT TCA	1408
Gly Gln Met Ser Gly Ala Arg Tyr Gly Gly Ser Ser Asn Ile Ala Ser	
430 435 440	
TTG ACC CCT GGG CCA GGC ATG CAA TCA CCA TCT TCC TAC CAG AAC AAC	1456
Leu Thr Pro Gly Pro Gly Met Gln Ser Pro Ser Ser Tyr Gln Asn Asn	
445 450 455	
AAC TAT GGG CTC AAC ATG AGT AGC CCC CCA CAT GGG AGT CCT GGT CTT	1504
Asn Tyr Gly Leu Asn Met Ser Ser Pro Pro His Gly Ser Pro Gly Leu	
460 465 470	
GCC CCA AAC CAG CAG AAT ATC ATG ATT TCT CCT CGT AAT CGT GGG AGT	1552
Ala Pro Asn Gln Gln Asn Ile Met Ile Ser Pro Arg Asn Arg Gly Ser	
475 480 485	
CCA AAG ATA GCC TCA CAT CAG TTT TCT CCT GTT GCA GGT GTG CAC TCT	1600
Pro Lys Ile Ala Ser His Gln Phe Ser Pro Val Ala Gly Val His Ser	
490 495 500 505	
CCC ATG GCA TCT TCT GGC AAT ACT GGG AAC CAC AGC TTT TCC AGC AGC	1648
Pro Met Ala Ser Ser Gly Asn Thr Gly Asn His Ser Phe Ser Ser Ser	
510 515 520	
TCT CTC AGT GCC CTG CAA GCC ATC AGT GAA GGT GTG GGG ACT TCC CTT	1696
Ser Leu Ser Ala Leu Gln Ala Ile Ser Glu Gly Val Gly Thr Ser Leu	
525 530 535	
TTA TCT ACT CTG TCA TCA CCA GGC CCC AAA TTG GAT AAC TCT CCC AAT	1744
Leu Ser Thr Leu Ser Ser Pro Gly Pro Lys Leu Asp Asn Ser Pro Asn	
540 545 550	
ATG AAT ATT ACC CAA CCA AGT AAA GTA AGC AAT CAG GAT TCC AAG AGT	1792
Met Asn Ile Thr Gln Pro Ser Lys Val Ser Asn Gln Asp Ser Lys Ser	
555 560 565	
CCT CTG GGC TTT TAT TGC GAC CAA AAT CCA GTG GAG AGT TCA ATG TGT	1840
Pro Leu Gly Phe Tyr Cys Asp Gln Asn Pro Val Glu Ser Ser Met Cys	
570 575 580 585	
CAG TCA AAT AGC AGA GAT CAC CTC AGT GAC AAA GAA AGT AAG GAG AGC	1888
Gln Ser Asn Ser Arg Asp His Leu Ser Asp Lys Glu Ser Lys Glu Ser	
590 595 600	
AGT GTT GAG GGG GCA GAG AAT CAA AGG GGT CCT TTG GAA AGC AAA GGT	1936
Ser Val Glu Gly Ala Glu Asn Gln Arg Gly Pro Leu Glu Ser Lys Gly	

605	610	615	
CAT AAA AAA TTA CTG CAG TTA CTT ACC TGT TCT TCT GAT GAC CGG GGT His Lys Lys Leu Leu Gln Leu Leu Thr Cys Ser Ser Asp Asp Arg Gly 620 625 630			1984
CAT TCC TCC TTG ACC AAC TCC CCC CTA GAT TCA AGT TGT AAA GAA TCT His Ser Ser Leu Thr Asn Ser Pro Leu Asp Ser Ser Cys Lys Glu Ser 635 640 645			2032
TCT GTT AGT GTC ACC AGC CCC TCT GGA GTC TCC TCC TCT ACA TCT GGA Ser Val Ser Val Thr Ser Pro Ser Gly Val Ser Ser Ser Thr Ser Gly 650 655 660 665			2080
GGA GTA TCC TCT ACA TCC AAT ATG CAT GGG TCA CTG TTA CAA GAG AAG Gly Val Ser Ser Thr Ser Asn Met His Gly Ser Leu Leu Gln Glu Lys 670 675 680			2128
CAC CGG ATT TTG CAC AAG TTG CTG CAG AAT GGG AAT TCA CCA GCT GAG His Arg Ile Leu His Lys Leu Leu Gln Asn Gly Asn Ser Pro Ala Glu 685 690 695			2176
GTA GCC AAG ATT ACT GCA CAA GCC ACT GGG AAA GAC ACC AGC AGT ATA Val Ala Lys Ile Thr Ala Gln Ala Thr Gly Lys Asp Thr Ser Ser Ile 700 705 710			2224
ACT TCT TGT GGG GAC GGA AAT GTT GTC AAG CAG GAG CAG CTA AGT CCT Thr Ser Cys Gly Asp Gly Asn Val Val Lys Gln Glu Gln Leu Ser Pro 715 720 725			2272
AAG AAG AAG GAG AAT AAT GCA CTT CTT AGA TAC CTG CTG GAC AGG GAT Lys Lys Lys Glu Asn Asn Ala Leu Leu Arg Tyr Leu Leu Asp Arg Asp 730 735 740 745			2320
GAT CCT AGT GAT GCA CTC TCT AAA GAA CTA CAG CCC CAA GTG GAA GGA Asp Pro Ser Asp Ala Leu Ser Lys Glu Leu Gln Pro Gln Val Glu Gly 750 755 760			2368
GTG GAC AAT AAA ATG AGT CAG TGC ACC AGC TCC ACC ATT CCT AGC TCA Val Asp Asn Lys Met Ser Gln Cys Thr Ser Ser Thr Ile Pro Ser Ser 765 770 775			2416
AGT CAA GAG AAA GAC CCT AAA ATT AAG ACA GAG ACA AGT GAA GAG GGA Ser Gln Glu Lys Asp Pro Lys Ile Lys Thr Glu Thr Ser Glu Glu Gly 780 785 790			2464
TCT GGA GAC TTG GAT AAT CTA GAT GCT ATT CTT GGT GAT CTG ACT AGT Ser Gly Asp Leu Asp Asn Leu Asp Ala Ile Leu Gly Asp Leu Thr Ser 795 800 805			2512
TCT GAC TTT TAC AAT AAT TCC ATA TCC TCA AAT GGT AGT CAT CTG GGG Ser Asp Phe Tyr Asn Asn Ser Ile Ser Ser Asn Gly Ser His Leu Gly 810 815 820 825			2560

ACT AAG CAA CAG GTG TTT CAA GGA ACT AAT TCT CTG GGT TTG AAA AGT	2608
Thr Lys Gln Gln Val Phe Gln Gly Thr Asn Ser Leu Gly Leu Lys Ser	
830 835 840	
TCA CAG TCT GTG CAG TCT ATT CGT CCT CCA TAT AAC CGA GCA GTG TCT	2656
Ser Gln Ser Val Gln Ser Ile Arg Pro Pro Tyr Asn Arg Ala Val Ser	
845 850 855	
CTG GAT AGC CCT GTT TCT GTT GGC TCA AGT CCT CCA GTA AAA AAT ATC	2704
Leu Asp Ser Pro Val Ser Val Gly Ser Ser Pro Pro Val Lys Asn Ile	
860 865 870	
AGT GCT TTC CCC ATG TTA CCA AAG CAA CCC ATG TTG GGT GGG AAT CCA	2752
Ser Ala Phe Pro Met Leu Pro Lys Gln Pro Met Leu Gly Gly Asn Pro	
875 880 885	
AGA ATG ATG GAT AGT CAG GAA AAT TAT GGC TCA AGT ATG GGT GGG CCA	2800
Arg Met Met Asp Ser Gln Glu Asn Tyr Gly Ser Ser Met Gly Gly Pro	
890 895 900 905	
AAC CGA AAT GTG ACT GTG ACT CAG ACT CCT TCC TCA GGA GAC TGG GGC	2848
Asn Arg Asn Val Thr Val Thr Gln Thr Pro Ser Ser Gly Asp Trp Gly	
910 915 920	
TTA CCA AAC TCA AAG GCC GGC AGA ATG GAA CCT ATG AAT TCA AAC TCC	2896
Leu Pro Asn Ser Lys Ala Gly Arg Met Glu Pro Met Asn Ser Asn Ser	
925 930 935	
ATG GGA AGA CCA GGA GGA GAT TAT AAT ACT TCT TTA CCC AGA CCT GCA	2944
Met Gly Arg Pro Gly Gly Asp Tyr Asn Thr Ser Leu Pro Arg Pro Ala	
940 945 950	
CTG GGT GGC TCT ATT CCC ACA TTG CCT CTT CGG TCT AAT AGC ATA CCA	2992
Leu Gly Gly Ser Ile Pro Thr Leu Pro Leu Arg Ser Asn Ser Ile Pro	
955 960 965	
GGT GCG AGA CCA GTA TTG CAA CAG CAG CAG CAG ATG CTT CAA ATG AGG	3040
Gly Ala Arg Pro Val Leu Gln Gln Gln Gln Gln Met Leu Gln Met Arg	
970 975 980 985	
CCT GGT GAA ATC CCC ATG GGA ATG GGG GCT AAT CCC TAT GGC CAA GCA	3088
Pro Gly Glu Ile Pro Met Gly Met Gly Ala Asn Pro Tyr Gly Gln Ala	
990 995 1000	
GCA GCA TCT AAC CAA CTG GGT TCC TGG CCC GAT GGC ATG TTG TCC ATG	3136
Ala Ala Ser Asn Gln Leu Gly Ser Trp Pro Asp Gly Met Leu Ser Met	
1005 1010 1015	
GAA CAA GTT TCT CAT GGC ACT CAA AAT AGG CCT CTT CTT AGG AAT TCC	3184
Glu Gln Val Ser His Gly Thr Gln Asn Arg Pro Leu Leu Arg Asn Ser	
1020 1025 1030	
CTG GAT GAT CTT GTT GGG CCA CCT TCC AAC CTG GAA GGC CAG AGT GAC	3232
Leu Asp Asp Leu Val Gly Pro Pro Ser Asn Leu Glu Gly Gln Ser Asp	

1035	1040	1045	
GAA AGA GCA TTA TTG GAC CAG CTG CAC ACT CTT CTC AGC AAC ACA GAT Glu Arg Ala Leu Leu Asp Gln Leu His Thr Leu Leu Ser Asn Thr Asp 1050 1055 1060 1065			3280
GCG ACA GGC CTG GAA GAA ATT GAC AGA GCT TTG GGC ATT CCT GAA CTT Ala Thr Gly Leu Glu Glu Ile Asp Arg Ala Leu Gly Ile Pro Glu Leu 1070 1075 1080			3328
GTC AAT CAG GGA CAG GCA TTA GAG CCC AAA CAG GAT GCT TTC CAA GGC Val Asn Gln Gly Gln Ala Leu Glu Pro Lys Gln Asp Ala Phe Gln Gly 1085 1090 1095			3376
CAA GAA GCA GCA GTA ATG ATG GAT CAG AAG GCA GGA TTA TAT GGA CAG Gln Glu Ala Ala Val Met Met Asp Gln Lys Ala Gly Leu Tyr Gly Gln 1100 1105 1110			3424
ACA TAC CCA GCA CAG GGG CCT CCA ATG CAA GGA GGC TTT CAT CTT CAG Thr Tyr Pro Ala Gln Gly Pro Pro Met Gln Gly Gly Phe His Leu Gln 1115 1120 1125			3472
GGA CAA TCA CCA TCT TTT AAC TCT ATG ATG AAT CAG ATG AAC CAG CAA Gly Gln Ser Pro Ser Phe Asn Ser Met Met Asn Gln Met Asn Gln Gln 1130 1135 1140 1145			3520
GGC AAT TTT CCT CTC CAA GGA ATG CAC CCA CGA GCC AAC ATC ATG AGA Gly Asn Phe Pro Leu Gln Gly Met His Pro Arg Ala Asn Ile Met Arg 1150 1155 1160			3568
CCC CGG ACA AAC ACC CCC AAG CAA CTT AGA ATG CAG CTT CAG CAG AGG Pro Arg Thr Asn Thr Pro Lys Gln Leu Arg Met Gln Leu Gln Gln Arg 1165 1170 1175			3616
CTG CAG GGC CAG CAG TTT TTG AAT CAG AGC CGA CAG GCA CTT GAA TTG Leu Gln Gly Gln Gln Phe Leu Asn Gln Ser Arg Gln Ala Leu Glu Leu 1180 1185 1190			3664
AAA ATG GAA AAC CCT ACT GCT GGT GGT GCT GCG GTG ATG AGG CCT ATG Lys Met Glu Asn Pro Thr Ala Gly Gly Ala Ala Val Met Arg Pro Met 1195 1200 1205			3712
ATG CAG CCC CAG CAG GGT TTT CTT AAT GCT CAA ATG GTC GCC CAA CGC Met Gln Pro Gln Gln Gly Phe Leu Asn Ala Gln Met Val Ala Gln Arg 1210 1215 1220 1225			3760
AGC AGA GAG CTG CTA AGT CAT CAC TTC CGA CAA CAG AGG GTG GCT ATG Ser Arg Glu Leu Leu Ser His His Phe Arg Gln Gln Arg Val Ala Met 1230 1235 1240			3808
ATG ATG CAG CAG CAG CAA CAG CAG CAG CAG CAG CAG CAG CAG CAG CAA Met Met Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln 1245 1250 1255			3856

CAG CAA CAG CAA CAG CAG CAA CAG CAG CAA ACC CAG GCC TTC AGC CCA Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Thr Gln Ala Phe Ser Pro 1260 1265 1270	3904
CCT CCT AAT GTG ACT GCT TCC CCC AGC ATG GAT GGG CTT TTG GCA GGA Pro Pro Asn Val Thr Ala Ser Pro Ser Met Asp Gly Leu Leu Ala Gly 1275 1280 1285	3952
CCC ACA ATG CCA CAA GCT CCT CCG CAA CAG TTT CCA TAT CAA CCA AAT Pro Thr Met Pro Gln Ala Pro Pro Gln Gln Phe Pro Tyr Gln Pro Asn 1290 1295 1300 1305	4000
TAT GGA ATG GGA CAA CAA CCA GAT CCA GCC TTT GGT CGA GTG TCT AGT Tyr Gly Met Gly Gln Gln Pro Asp Pro Ala Phe Gly Arg Val Ser Ser 1310 1315 1320	4048
CCT CCC AAT GCA ATG ATG TCG TCA AGA ATG GGT CCC TCC CAG AAT CCC Pro Pro Asn Ala Met Met Ser Ser Arg Met Gly Pro Ser Gln Asn Pro 1325 1330 1335	4096
ATG ATG CAA CAC CCG CAG GCT GCA TCC ATC TAT CAG TCC TCA GAA ATG Met Met Gln His Pro Gln Ala Ala Ser Ile Tyr Gln Ser Ser Glu Met 1340 1345 1350	4144
AAG GGC TGG CCA TCA GGA AAT TTG GCC AGG AAC AGC TCC TTT TCC CAG Lys Gly Trp Pro Ser Gly Asn Leu Ala Arg Asn Ser Ser Phe Ser Gln 1355 1360 1365	4192
CAG CAG TTT GCC CAC CAG GGG AAT CCT GCA GTG TAT AGT ATG GTG CAC Gln Gln Phe Ala His Gln Gly Asn Pro Ala Val Tyr Ser Met Val His 1370 1375 1380 1385	4240
ATG AAT GGC AGC AGT GGT CAC ATG GGA CAG ATG AAC ATG AAC CCC ATG Met Asn Gly Ser Ser Gly His Met Gly Gln Met Asn Met Asn Pro Met 1390 1395 1400	4288
CCC ATG TCT GGC ATG CCT ATG GGT CCT GAT CAG AAA TAC TGC TGA CAT CT Pro Met Ser Gly Met Pro Met Gly Pro Asp Gln Lys Tyr Cys * 1405 1410 1415	4338
CTGCACCAGG ACCTCTTAAG GAAACCACTG TACAAATGAC ACTGCACTAG GATTATTGGG	4398
AAGGAATCAT TGTTCAGGC ATCCATCTTG GAAGAAAGGA CCAGCTTTGA GCTCCATCAA	4458
GGGTATTTTA AGTGATGTCA TTTGAGCAGG AATTCTAG	4496

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1417 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met	Ser	Gly	Leu	Gly	Glu	Asn	Leu	Asp	Pro	Leu	Ala	Ser	Asp	Ser	Arg	1	5	10	15
Lys	Arg	Lys	Leu	Pro	Cys	Asp	Thr	Pro	Gly	Gln	Gly	Leu	Thr	Cys	Ser	20	25	30	
Gly	Glu	Lys	Arg	Arg	Arg	Glu	Gln	Glu	Ser	Lys	Tyr	Ile	Glu	Glu	Leu	35	40	45	
Ala	Glu	Leu	Ile	Ser	Ala	Asn	Leu	Ser	Asp	Ile	Asp	Asn	Phe	Asn	Val	50	55	60	
Lys	Pro	Asp	Lys	Cys	Ala	Ile	Leu	Lys	Glu	Thr	Val	Arg	Gln	Ile	Arg	65	70	75	80
Gln	Ile	Lys	Glu	Gln	Gly	Lys	Thr	Ile	Ser	Asn	Asp	Asp	Asp	Val	Gln	85	90	95	
Lys	Ala	Asp	Val	Ser	Ser	Thr	Gly	Gln	Gly	Val	Ile	Asp	Lys	Asp	Ser	100	105	110	
Leu	Gly	Pro	Leu	Leu	Leu	Gln	Ala	Leu	Asp	Gly	Phe	Leu	Phe	Val	Val	115	120	125	
Asn	Arg	Glu	Ala	Asn	Ile	Val	Phe	Val	Ser	Glu	Asn	Val	Thr	Gln	Tyr	130	135	140	
Leu	Gln	Tyr	Lys	Gln	Glu	Asp	Leu	Val	Asn	Thr	Ser	Val	Tyr	Asn	Ile	145	150	155	160
Leu	His	Glu	Glu	Asp	Arg	Lys	Asp	Phe	Leu	Lys	Asn	Leu	Pro	Lys	Ser	165	170	175	
Thr	Val	Asn	Gly	Val	Ser	Trp	Thr	Asn	Glu	Pro	Gln	Arg	Gln	Lys	Ser	180	185	190	
His	Thr	Phe	Asn	Cys	Arg	Met	Leu	Met	Lys	Thr	Pro	His	Asp	Ile	Leu	195	200	205	
Glu	Asp	Ile	Asn	Ala	Ser	Pro	Glu	Met	Arg	Gln	Arg	Tyr	Glu	Thr	Met	210	215	220	
Gln	Cys	Phe	Ala	Leu	Ser	Gln	Pro	Arg	Ala	Met	Met	Glu	Glu	Gly	Glu	225	230	235	240
Asp	Leu	Gln	Ser	Cys	Met	Ile	Cys	Val	Ala	Arg	Arg	Ile	Thr	Thr	Gly	245	250	255	
Glu	Arg	Thr	Phe	Pro	Ser	Asn	Pro	Glu	Ser	Phe	Ile	Thr	Arg	His	Asp	260	265	270	

Leu Ser Gly Lys Val Val Asn Ile Asp Thr Asn Ser Leu Arg Ser Ser
 275 280 285
 Met Arg Pro Gly Phe Glu Asp Ile Ile Arg Arg Cys Ile Gln Arg Phe
 290 295 300
 Phe Ser Leu Asn Asp Gly Gln Ser Trp Ser Gln Lys Arg His Tyr Gln
 305 310 315 320
 Glu Ala Tyr Leu Asn Gly His Ala Glu Thr Pro Val Tyr Arg Phe Ser
 325 330 335
 Leu Ala Asp Gly Thr Ile Val Thr Ala Gln Thr Lys Ser Lys Leu Phe
 340 345 350
 Arg Asn Pro Val Thr Asn Asp Arg His Gly Phe Val Ser Thr His Phe
 355 360 365
 Leu Gln Arg Glu Gln Asn Gly Tyr Arg Pro Asn Pro Asn Pro Val Gly
 370 375 380
 Gln Gly Ile Arg Pro Pro Met Ala Gly Cys Asn Ser Ser Val Gly Gly
 385 390 395 400
 Met Ser Met Ser Pro Asn Gln Gly Leu Gln Met Pro Ser Ser Arg Ala
 405 410 415
 Tyr Gly Leu Ala Asp Pro Ser Thr Thr Gly Gln Met Ser Gly Ala Arg
 420 425 430
 Tyr Gly Gly Ser Ser Asn Ile Ala Ser Leu Thr Pro Gly Pro Gly Met
 435 440 445
 Gln Ser Pro Ser Ser Tyr Gln Asn Asn Asn Tyr Gly Leu Asn Met Ser
 450 455 460
 Ser Pro Pro His Gly Ser Pro Gly Leu Ala Pro Asn Gln Gln Asn Ile
 465 470 475 480
 Met Ile Ser Pro Arg Asn Arg Gly Ser Pro Lys Ile Ala Ser His Gln
 485 490 495
 Phe Ser Pro Val Ala Gly Val His Ser Pro Met Ala Ser Ser Gly Asn
 500 505 510
 Thr Gly Asn His Ser Phe Ser Ser Ser Ser Leu Ser Ala Leu Gln Ala
 515 520 525
 Ile Ser Glu Gly Val Gly Thr Ser Leu Leu Ser Thr Leu Ser Ser Pro
 530 535 540
 Gly Pro Lys Leu Asp Asn Ser Pro Asn Met Asn Ile Thr Gln Pro Ser
 545 550 555 560

Lys Val Ser Asn Gln Asp Ser Lys Ser Pro Leu Gly Phe Tyr Cys Asp
 565 570 575
 Gln Asn Pro Val Glu Ser Ser Met Cys Gln Ser Asn Ser Arg Asp His
 580 585 590
 Leu Ser Asp Lys Glu Ser Lys Glu Ser Ser Val Glu Gly Ala Glu Asn
 595 600 605
 Gln Arg Gly Pro Leu Glu Ser Lys Gly His Lys Lys Leu Leu Gln Leu
 610 615 620
 Leu Thr Cys Ser Ser Asp Asp Arg Gly His Ser Ser Leu Thr Asn Ser
 625 630 635 640
 Pro Leu Asp Ser Ser Cys Lys Glu Ser Ser Val Ser Val Thr Ser Pro
 645 650 655
 Ser Gly Val Ser Ser Ser Thr Ser Gly Gly Val Ser Ser Thr Ser Asn
 660 665 670
 Met His Gly Ser Leu Leu Gln Glu Lys His Arg Ile Leu His Lys Leu
 675 680 685
 Leu Gln Asn Gly Asn Ser Pro Ala Glu Val Ala Lys Ile Thr Ala Gln
 690 695 700
 Ala Thr Gly Lys Asp Thr Ser Ser Ile Thr Ser Cys Gly Asp Gly Asn
 705 710 715 720
 Val Val Lys Gln Glu Gln Leu Ser Pro Lys Lys Lys Glu Asn Asn Ala
 725 730 735
 Leu Leu Arg Tyr Leu Leu Asp Arg Asp Asp Pro Ser Asp Ala Leu Ser
 740 745 750
 Lys Glu Leu Gln Pro Gln Val Glu Gly Val Asp Asn Lys Met Ser Gln
 755 760 765
 Cys Thr Ser Ser Thr Ile Pro Ser Ser Ser Gln Glu Lys Asp Pro Lys
 770 775 780
 Ile Lys Thr Glu Thr Ser Glu Glu Gly Ser Gly Asp Leu Asp Asn Leu
 785 790 795 800
 Asp Ala Ile Leu Gly Asp Leu Thr Ser Ser Asp Phe Tyr Asn Asn Ser
 805 810 815
 Ile Ser Ser Asn Gly Ser His Leu Gly Thr Lys Gln Gln Val Phe Gln
 820 825 830
 Gly Thr Asn Ser Leu Gly Leu Lys Ser Ser Gln Ser Val Gln Ser Ile
 835 840 845

Arg Pro Pro Tyr Asn Arg Ala Val Ser Leu Asp Ser Pro Val Ser Val
 850 855 860
 Gly Ser Ser Pro Pro Val Lys Asn Ile Ser Ala Phe Pro Met Leu Pro
 865 870 875 880
 Lys Gln Pro Met Leu Gly Gly Asn Pro Arg Met Met Asp Ser Gln Glu
 885 890 895
 Asn Tyr Gly Ser Ser Met Gly Gly Pro Asn Arg Asn Val Thr Val Thr
 900 905 910
 Gln Thr Pro Ser Ser Gly Asp Trp Gly Leu Pro Asn Ser Lys Ala Gly
 915 920 925
 Arg Met Glu Pro Met Asn Ser Asn Ser Met Gly Arg Pro Gly Gly Asp
 930 935 940
 Tyr Asn Thr Ser Leu Pro Arg Pro Ala Leu Gly Gly Ser Ile Pro Thr
 945 950 955 960
 Leu Pro Leu Arg Ser Asn Ser Ile Pro Gly Ala Arg Pro Val Leu Gln
 965 970 975
 Gln Gln Gln Gln Met Leu Gln Met Arg Pro Gly Glu Ile Pro Met Gly
 980 985 990
 Met Gly Ala Asn Pro Tyr Gly Gln Ala Ala Ala Ser Asn Gln Leu Gly
 995 1000 1005
 Ser Trp Pro Asp Gly Met Leu Ser Met Glu Gln Val Ser His Gly Thr
 1010 1015 1020
 Gln Asn Arg Pro Leu Leu Arg Asn Ser Leu Asp Asp Leu Val Gly Pro
 1025 1030 1035 1040
 Pro Ser Asn Leu Glu Gly Gln Ser Asp Glu Arg Ala Leu Leu Asp Gln
 1045 1050 1055
 Leu His Thr Leu Leu Ser Asn Thr Asp Ala Thr Gly Leu Glu Glu Ile
 1060 1065 1070
 Asp Arg Ala Leu Gly Ile Pro Glu Leu Val Asn Gln Gly Gln Ala Leu
 1075 1080 1085
 Glu Pro Lys Gln Asp Ala Phe Gln Gly Gln Glu Ala Ala Val Met Met
 1090 1095 1100
 Asp Gln Lys Ala Gly Leu Tyr Gly Gln Thr Tyr Pro Ala Gln Gly Pro
 1105 1110 1115 1120
 Pro Met Gln Gly Gly Phe His Leu Gln Gly Gln Ser Pro Ser Phe Asn
 1125 1130 1135

Ser Met Met Asn Gln Met Asn Gln Gln Gly Asn Phe Pro Leu Gln Gly
 1140 1145 1150
 Met His Pro Arg Ala Asn Ile Met Arg Pro Arg Thr Asn Thr Pro Lys
 1155 1160 1165
 Gln Leu Arg Met Gln Leu Gln Gln Arg Leu Gln Gly Gln Gln Phe Leu
 1170 1175 1180
 Asn Gln Ser Arg Gln Ala Leu Glu Leu Lys Met Glu Asn Pro Thr Ala
 1185 1190 1195 1200
 Gly Gly Ala Ala Val Met Arg Pro Met Met Gln Pro Gln Gln Gly Phe
 1205 1210 1215
 Leu Asn Ala Gln Met Val Ala Gln Arg Ser Arg Glu Leu Leu Ser His
 1220 1225 1230
 His Phe Arg Gln Gln Arg Val Ala Met Met Met Gln Gln Gln Gln Gln
 1235 1240 1245
 Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln
 1250 1255 1260
 Gln Gln Gln Thr Gln Ala Phe Ser Pro Pro Pro Asn Val Thr Ala Ser
 1265 1270 1275 1280
 Pro Ser Met Asp Gly Leu Leu Ala Gly Pro Thr Met Pro Gln Ala Pro
 1285 1290 1295
 Pro Gln Gln Phe Pro Tyr Gln Pro Asn Tyr Gly Met Gly Gln Gln Pro
 1300 1305 1310
 Asp Pro Ala Phe Gly Arg Val Ser Ser Pro Pro Asn Ala Met Met Ser
 1315 1320 1325
 Ser Arg Met Gly Pro Ser Gln Asn Pro Met Met Gln His Pro Gln Ala
 1330 1335 1340
 Ala Ser Ile Tyr Gln Ser Ser Glu Met Lys Gly Trp Pro Ser Gly Asn
 1345 1350 1355 1360
 Leu Ala Arg Asn Ser Ser Phe Ser Gln Gln Gln Phe Ala His Gln Gly
 1365 1370 1375
 Asn Pro Ala Val Tyr Ser Met Val His Met Asn Gly Ser Ser Gly His
 1380 1385 1390
 Met Gly Gln Met Asn Met Asn Pro Met Pro Met Ser Gly Met Pro Met
 1395 1400 1405
 Gly Pro Asp Gln Lys Tyr Cys *
 1410 1415